The Pennsylvania State University

The Graduate School

College of Education

TALK ALOUD PROBLEM SOLVING: EXPLORATION OF ACQUISITION AND
FREQUENCY BUILDING IN SCIENCE TEXT

A Dissertation in

Special Education

by

Ginny Dembek

© 2013 Ginny Dembek

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2013
The dissertation of Ginny A. Dembek was reviewed and approved* by the following:

Richard M. Kubina Jr.
Professor of Education
Dissertation Advisor
Chair of Committee
Department Head

David L. Lee
Associate Professor of Education

Mary Catherine Scheeler
Associate Professor of Education

Scott McDonald
Associate Professor of Education

*Signatures are on file in the Graduate School
Abstract

Discovering new ways to help students attain higher levels of scientific knowledge and to think critically is a national goal (Educate to Innovate campaign). Despite the best intentions, many students struggle to achieve a basic level of science knowledge (NAEP, 2011). The present study examined Talk Aloud Pair Problem Solving and frequency building with five students who were diagnosed with a disability and receive specialized reading instruction in a special education setting. Acquisition was obtained through scripted lessons and frequency building or practice strengthened the student’s verbal repertoire making the problem solving process a durable behavior. Overall, students all demonstrated improvements in problem solving performance when compared to baseline. Students became more significantly accurate in performance and maintenance in learning was demonstrated. Generalization probes indicated improvement in student performance. Implications for practice and future research are discussed.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... v

Acknowledgements ..................................................................................................... vi

Introduction ................................................................................................................ 1

  Research Questions .......................................................................................... 6

Method ...................................................................................................................... 7

  Participants ........................................................................................................ 7
  Setting .................................................................................................................. 9
  Materials ............................................................................................................. 10
  Data Collection ................................................................................................. 10
  Experimental Design .......................................................................................... 12

Results ...................................................................................................................... 19

  Frequency measures ....................................................................................... 20
  Maintenance ...................................................................................................... 20
  Generalization probes .................................................................................... 24
  Social Validity ................................................................................................... 26

Discussion .................................................................................................................. 27

  Frequency measures ....................................................................................... 28
  Maintenance ...................................................................................................... 32
  Generalization probes .................................................................................... 32

Future directions and Limitations ........................................................................... 35

References ............................................................................................................... 37

Figure 1 .................................................................................................................. 44

Figure 2 .................................................................................................................. 45

Appendices

  Appendix A: Review of the Literature.............................................................. 46
  Appendix B: Instructional Material ................................................................. 74
  Appendix C: Dependent Variable Sample ....................................................... 81
  Appendix D: Dependent Variable Answer Key for Scoring ......................... 82
  Appendix E: Independent Variable Transcripts and Scoring ....................... 84
LIST OF FIGURES

Figure 1: All Participants Results for Sequence of Sentences per 2-Minute Across Phases .................................................................44

Figures 2: All Students Participants Results for Generalization Sequence of Sentences per 2-Minute Across Phases .........................................................45
Acknowledgements

First and foremost, I would like to thank my advisor and mentor, Dr. Richard Kubina, for his guidance, encouragement, and advice. I have been extremely lucky to have a professor who helped me explore my research interests and shared in my excitement about future endeavors. I would not have had the opportunity to complete my research, had it not been for his support and encouragement. The discussion and experiences he shared will forever shape the type of educator and researcher I strive to be.

I would also like to thank my doctoral committee. I appreciate the experience and knowledge they contribute to my research. They have challenged me to always become a better researcher and contributor to the field. Additionally, I would like to thank all the people at Penn State who have encouraged me during my time in undergraduate and graduate studies.

I must express gratitude to my parents, who have always been my number one supporters. They helped shape my belief in what great educators should be through their own actions. Along with my parents I need to thank my amazing family, especially my brothers, for always being there for me. Having such a strong support system throughout my life has guided me to my current path and for that I couldn’t be more thankful.

Finally, I want to thank my loving husband for his endless support. He never questioned my ability to succeed, encouraging me during difficult times. For the weeks I spent by his side and the months I spent far away, he never hesitated to support me.
Talk Aloud Problem Solving: Exploration of Acquisition and Frequency Building in Science Text

The world is constantly changing and evolving with new technologies and discoveries. Science and scientific literacy are becoming increasingly important as the United States competes with developing countries and engages in debates on decision making within the country (National Research Council, 1996, pp. 1-2). People are exposed to science issues in policy debates, decisions that could be informed by science, and the everyday use of science (Marincola, 2006). The sheer number of decisions influenced by scientific knowledge is astounding (e.g., the effect of smoking on the body, the effect of varied caloric intake, determining chemical additives in food, navigation of new software on tablet computers, or ide effects of medication on the body). A person who is not scientifically literate is automatically at a disadvantage because of the lack of information and skills necessary to engage thoughtfully in ongoing discussions as an informed, independent thinker (Marincola, 2006). The place to begin educating society is in the classroom.

Science is receiving the attention it deserves through mandates and other important legislation (e.g., President Obama’s Educate to Innovate Campaign). Despite the increased attention, a large segment of students have trouble mastering even the basics of science. A lack of understanding makes information that scientists share not clear, resulting in confusion for otherwise easy solutions (e.g., not understanding how to lose weight by reducing calories). When it comes to a group with a significant problem, students with disabilities have greater difficulties compared to their peers with 66% performing below basic according to the most recent NAEP Science results (NAEP, 2011). The number is more than twice the rate of students
without disabilities who are English speakers (NAEP, 2011). Science is difficult because it contains complex expository text compared with narrative text (Saenz & Fuchs, 2002). The majority of students in the learning disability category are included for their content area subjects with general education instruction.

Classes such as science can provide a significant challenge for students with learning disabilities because reading is where many already need more intense instruction. Students with learning disabilities demonstrate more severe forms of reading problems in comparison to poor readers not identified as having a learning disability (Fuchs, Fuchs, Mathes, & Lipsey, 2000). As the reading materials become more difficult, especially in content areas like science, growing numbers of students need more intense and effective instruction to fully benefit from both in-class and out-of-class activities.

The combination of legislation in the Individuals with Disabilities Education Act (IDEA, 2004) and No Child Left Behind (NCLB, 2001) created an expectation that science is an area that children with disabilities (i.e., with appropriate accommodations) reach the same level as their peers in relation to general education curriculum standards. Because students are included and expected to learn the same content, research needs to continue to explore what is currently present in the classroom and how it can be modified (Brigham, Scruggs, & Mastropieri, 2011). Compounding the difficulty teachers face finding the best way to teach content to all students, little is published in the field of learning disabilities research in science. A recent analysis of 11 major journals found that only 15.9 percent of articles were intervention research with only seven percent of those representing content area intervention research over a 19 year period (i.e., including social studies research) (Mastropieri et al., 2009). The need is apparent for more
research if students with disabilities are expected to learn the same content as their general education peers.

Expository text introduces many challenges and complexities increasing difficulties for content area learning. Students with disabilities, for example, receiving intense reading instruction as per their Individualized Education Plan (IEP) and would still be expected to perform with some minor accommodations to solve problems in content areas such as science. Expository text brings additional challenges with advanced vocabulary, varied and complex text structure, conceptual density, and a requirement for additional prior knowledge (Saenz & Fuchs, 2002). The difficulty of comprehending more complex types of text, however, can be lessened by teaching students to problem solve. One intervention that has shown promise is Talk Aloud Problem Solving or TAPS (Whimbey & Lochhead, 1999). The structured TAPS intervention includes modeling and prompting of think aloud statements and can be implemented across skills and content areas. Research with talk alouds, also called think alouds, during reading has demonstrated the importance of including explicit modeling and prompting with instruction (Baumann, Seifert-Kessell, & Jones, 1992; Bereiter & Bird, 1985; Witcoski, 2012).

Within science instruction, the literature is vast covering concepts from cell mitosis and chemical reactions to the water cycle and evolution. There have been different individual strategies implemented for problem solving instruction in science such as group brainstorming (DeHann, 2009). Unfortunately, group strategies for problem solving are limited when others for appropriate pairing or selection are not present. In comparison, TAPS is designed specifically for an individual student and can be implemented with pairs of students as they master the
process of TAPS. Most importantly, the student benefits by becoming his or her own listener and more independent (Robbins, 2011).

TAPS has been used in high school science class (Pestel, 1993) including chemistry (Tingle & Good, 1990), mechanics instruction with a multimedia component (Holzer & Anduret, 2000), and troubleshooting in aviation technician university instruction (Johnson & Chung, 1999). The transfer to the application of solving problems in science (e.g., equations in chemistry, experiments in physics) is directly relevant because students learn to break down skills into smaller components. While TAPS is a promising intervention, the majority of studies have been conducted with mostly high school and college age students.

A recent study extended the TAPS literature by implementing the problem solving strategy using science text in fourth grade. Fourth grade is the age students are come into contact with more expository compared to mostly narrative text (Grigg, Daane, Jin, & Campbell, 2003). Witcoski (2012) used TAPS with students in fourth grade at-risk for learning disabilities who demonstrated difficulties in reading comprehension and problem solving. The problem solving task of placing sentences from a science textbook (i.e., four sentences out of order) in a logical order was selected as the dependent variable. The task was chosen for its effect on students ability to organize a passage while demonstrating their awareness of text structure. The independent variable included TAPS instruction and frequency building (i.e., practice) talk aloud statements pertaining to decision making (i.e., talk aloud statements about cues in the sentences). The experimenter found that TAPS instruction and frequency building demonstrated a positive effect on the students ability to place sentences in a logical order.
The positive effect in the TAPS study reflected the student’s ability to problem solve and determine a logical order of content. The students used text structure, concepts, vocabulary, and prior knowledge to help them, all areas making science text a more difficult text to read (Saenz & Fuchs, 2002). Witcoski (2012) began to address a previous concern suggesting students with learning disabilities were not as aware of passage organization (i.e., text structure) and struggled with reorganizing disorganized passages when compared to students without learning disabilities (Wong & Wilson, 1984). Strategies helping students navigate text structure (e.g., TAPS) and apply appropriate structure-specific strategies may be beneficial across varied prose (Bakken, Mastropieri, & Scruggs, 1997) and science (Witcoski, 2012).

One way to engage the language of problem solving in science is through TAPS. TAPS has been implemented in high school and college settings with success. Witcoski (2012) demonstrated how TAPS can be useful with fourth grade students. Students who practiced the language of problem solving, to explain their process like older advanced students (i.e, students in advanced placement science courses in high school) stated more details in explanations. However, the fourth grade students were at-risk and not with an identified learning disability.

Therefore, to address the needs of students with learning disabilities struggling in the science classroom, applying TAPS may serve as an effective solution. Schools are required to gain information about student performance prior to the student failing or being placed in a more restrictive environment. Response to Intervention (RTI) is the reaction to the need and requires constant monitoring of student performance as well as universal screening and research-based instruction. Specifically, RTI is an array of procedures that can be used to determine if and how students respond to changes in instruction (Klotz & Canter, 2007). In the Response to
Intervention framework, TAPS can be a measure to prevent students from requiring additional instruction in Tier 2 or Tier 3 (i.e., Tier 1 is general education class-wide instruction, Tier 2 is more specific small group interventions targeting student difficulties, and Tier 3 is individualized instruction usually in the form of special education) (Fuchs & Fuchs, 2007).

In Tier 3, students with learning disabilities would learn to vocalize their process of problem solving and self-advocate for additional instruction, clarification of text, or accommodations that may be needed. The TAPS intervention fosters a skill that students with learning disabilities can benefit from in science. To explore TAPS with students with learning disabilities in reading, the following questions were asked: (1) Does TAPS and talking aloud to a frequency criterion demonstrate an effect on students with learning disabilities using a problem solving skill involving science text? (2) Do students maintain the ability to solve problems over time following TAPS and frequency building. (3) Does the problem solving skill targeting science content generalize to other content areas, specifically social studies? (4) Do teachers and students view the intervention of TAPS and frequency building as a socially valid and beneficial classroom tool?
Method

Participants

The experimenter outlined specific procedures to identify participants who benefited from inclusion in the TAPS intervention study. First, teachers nominated students in their classroom that struggled with test-taking and reading comprehension skills. Parental consent was sent home to allow the experimenter to test and determine if students would benefit from the study. Following nomination and parental consent, each student was tested on current reading fluency and retell ability. The experimenter identified students who struggled with comprehension not due primarily to decoding fluency problems. Students with a low reading fluency, (less than the 25th percentile for winter on grade level) specifically below 87 correct words per minute (CWPM), could struggle to benefit from the intervention because of the need for more intense lessons in decoding. The oral fluency norms were obtained from Hasbrouck and Tindal (2006). Along with decoding fluency, students were asked to retell what they remembered after reading a one minute passage. The students needed to retell five or less thought units (i.e., independent clause) per passage in one minute based on half of a fluent retell performance (Culler, 2010). An example of a sentence that would count as two independent clauses, or two retells is the following. The car drove down the street/and she waved goodbye. The reading fluency and retell scores needed to occur in two of three passages in order for the student to qualify for the study.

Abbie was a 4th grade female Caucasian student who was diagnosed with a specific learning disability in reading and writing. When tested for grade level reading fluency using
DIBELS passages Abbie read 87 CWPM with five incorrect words per minute (IWPM) on passage one, 71 CWPM with six IWPM on passage two, and 52 WCPM with six IWPM on passage three. Abbie was able to produce six retells on passage one, two correct and one incorrect retells on passage two, and two correct and one incorrect retells on passage three.

Jada was the second 4th grade female Caucasian student to begin the intervention. She was diagnosed with Other Health Impairments due to Attention Deficit Disorder and received instruction in the special education setting for reading, writing and mathematics. She completed tests for grade level reading fluency using the same DIBELS passages reading 101 CWPM with four IWPM on passage one, 79 CWPM with four IWPM on passage two, and 59 CWPM with five IWPM on passage three. When asked to retell information from the passages Jada was able to produce one correct retell on passage one, two and one-half correct retells on passage two, and one correct retell on passage three.

Saddie was a 4th grade female Caucasian student diagnosed with a specific learning disability in reading. When tested on grade level reading fluency using the same DIBELS passages she read 65 CWPM with four IWPM on passage one, 61 CWPM with six IWPM on passage two, and 57 CWPM with five IWPM on passage three. Her retell performance on the same passages was four correct retells on passage one, four correct retells on passage two, and three correct retells on passage three.

The fourth student was Sasha. Sasha was a 4th grade female African American student diagnosed with a specific learning disability in reading and other health impairment. She was tested on grade level reading fluency using the same DIBELS passages as other participants and read 89 CWPM with three IWPM on passage one, 79 CWPM with five IWPM on passage two,
and 88 CWPM with four IWPM on passage three. She was able to produce two and one-half correct retells for passage one, two and one-half correct retells for passage two, and two correct retells for passage three.

Theodore was a Caucasian male in 4th grade and was diagnosed with Other Health Impairment due to fetal alcohol syndrome and received instruction in reading, writing and mathematics in the learning support classroom. He was tested on grade level reading fluency using the same DIBELS passages and read 101 CWPM with three IWPM on passage one, 76 CWPM with three IWPM on passage two, 94 CWPM with one IWPM on passage three. On the same passages Theodore was able to retell two correct retells for passage one, three correct and one incorrect retells for passage two, three correct and one-half correct retells for passage three.

**Setting**

The study was conducted in a suburban/rural public school district located in the northeastern United States. The location of the delivery of the intervention was at the convenience of the teacher and students in two separate elementary schools. Typically instruction occurred in the learning support classroom or in the hallway directly outside of the classroom. Depending on student schedules, the student was pulled from the learning support classroom or general education classroom for the study. The majority of days, the intervention occurred during learning support instruction. The total enrollment of the school district for 2012-2013 school year was approximately 6,700 students. The school district provided special education to approximately 785 students with 20% of all students qualifying for free/reduced lunch. The school district demographic information shows 84% of students identifying as
Caucasian, 3% of each the following identifying as Hispanic, mixed race, and African American, 7 % identified as Asian.

**Dependent Variable**

Students were asked to answer a basic problem solving task referred to as “sequencing sentences in a logical order” related to reading comprehension (Whimbey, 1995). The task is presented as four sentences and the student must read each sentence and write the correct order. The order was written using a number from 1 to 4 next to the sentence which indicated the correct order of sequence. The dependent variable was timed for two minutes.

The materials used for the dependent variable, along with the independent variable, were collected from science textbooks currently used in classrooms (see Appendix C). Some materials had already been developed from textbooks for fourth grade students (Banks et al., 2001; Cooney et al., 2006; Dawson-Boyd et al., 2006; Hacket et al., 2008; Heil et. al., 1994; and Sciencesaurus, 2002) and used in a previous study (Witcoski, 2012). The sentences chosen included cues that helped students place them in order. The cues included transition words, noun phrases, meaning, events in a sequence, and wording (e.g., using a proper noun in the beginning sentence and a pronoun in a latter sentence).

Each page of the dependent variable contained four sets of problems (as shown in Appendix C). Each problem had four sentences that could be linked consecutively. The student had the possibility of 16 correct responses on each page, with another page available if needed. To complete the problems, the student indicated the order of the sentences by letter or number (i.e., taught number but some students may find using letters easier) in the spaces located in front of the sentences. To prevent a student from completing a large number of incorrect responses
without reading the prompt in baseline, a provision was put in place requiring the student to read the dependent variable aloud to guard against guessing.

The experimenter found sentences in textbooks and had two people (i.e., two students in undergraduate studies), check the prompts to ensure they could be solved correctly and that none were confusing. The experimenter threw out prompts, removing all the problems completed incorrectly or that were confusing to adult learners. The prompts collected from textbooks were located by finding a set of four sentences in the same paragraph (i.e., beginning sentence needed to start the paragraph). Prompts were tested and were able to be solved correctly without confusion by adult learners.

**Independent Variable**

The independent variable was based on Whimbey and Lochhead’s Thinking Aloud Problem Solving, the Talk Aloud Problem Solving (TAPS) method. The experimenter developed two stages for the intervention, the first coming from Whimbey and Lochhead (1999). The first part of the independent variable was the accuracy building stage where the experimenter modeled the process of talking aloud as shown in Appendix B. Following the model given through scripted lessons, the student solved problems with prompting and feedback from the expert (e.g., during the model the student only was shown the science text problem, not the entire script). The experimenter served as the expert and followed scripts to provide similar instruction across students. After each lesson, the student was given two minutes to talk aloud one problem (i.e., four sentences out of order). The student needed to read all of the sentences and talk aloud the problem and solution. Talking aloud was important to vocalize otherwise unknown information that the student was thinking while working. Following the two minute task to
talking aloud the correct sequence of sentences, the dependent variable, writing the correct sequence of sentences was given to the student for two minutes.

The second stage of the TAPS Independent Variable was the frequency building or the practice stage. The students continued to practice talking aloud during problem solving with immediate feedback. The student was timed performing the task of talking aloud the correct sequence of sentences for two minutes, with one minute of feedback from the expert (i.e., the experimenter). After feedback was given, the student repeated talking aloud the same problem for two minutes with another one minute of feedback. Similar to the first stage, the writing the correct sequence of sentences task was given to students following the second minute of feedback.

**Experimental Design**

The experimental design was a multiple baseline design across participants (Gast, 2010). Multiple baseline design across participants is useful for skills that cannot be unlearned, such as many academic skills (Cooper, Heron, & Heward, 2007). The multiple baseline design demonstrates strong internal validity if a functional relation is demonstrated between the independent and dependent variables in a study (Kennedy, 2005). All data are displayed on a Standard Celeration Chart (SCC).

**Accuracy of Dependent Variable**

Accuracy was determined to find the extent to which observed values estimated the events that took place in an experiment (Johnston & Pennypacker, 2009). The experimenter scored the dependent variable, writing the correct sequence of sentences. As an additional measure to score more accurately, the experimenter scored all of the problems twice. The second
scoring represents effort the experimenter took to ensure that error has been minimized in some way, also known as the true value (Johnston & Pennypacker, 2009). After scoring, an independent scorer re-scored 33 percent of the writing the correct sequence of sentences to check for accuracy. The scorer followed an answer key to check the answers (i.e., training was simply teaching the scorer how to use the answer key with problem codes). While re-scoring the dependent variable, the scorer checked off a separate sheet with the random numbers while re-scoring the probes (see Appendix C) indicating they checked the number and if it was scored correctly or incorrectly (i.e., wrote what the scoring should be and what it is).

The dependent variable was created along with an answer key containing the paragraphs along with the problem and answers. The original scorer, the experimenter, checked all probes by having undergraduates complete them (i.e., and correcting or throwing out problems that were incorrect). The observer was a fourth-year special education major and honors student. The independent observer went through the answer key and checked 33% randomly selected dependent variable probes demonstrating 99% accuracy in scoring.

**Procedural Integrity**

One independent scorer listened to 33 percent of the lessons and determined the procedural integrity. The scorer used the scripts (Appendix B) the experimenter implemented in the study and followed along with the audio version. When the scorer heard the content included in each paragraph, she checked off the area in front of the part of the lesson.

An additional measure to improve the procedural integrity is the scoring of the independent variable measure (i.e., talking aloud the correct sequence of sentences). The calibration, or evaluation of data from the measurement procedure, helped to adjust and improve
the procedure (Johnston & Pennypacker, 2009). The scorer wrote directly on the transcribed sheets (i.e., the talking aloud the correct order of sentences task) and separated the statements into the appropriate three categories. Training was required for the scorer in order to reach 100% agreement on at least three examples. A sample of re-scoring talking aloud the correct sequence of sentences scoring is attached in Appendix D. The sentences in the problem are highlighted yellow and the independent scorer is red. The talk alouds about cue (i.e., the measure that determined students entered the next stage or completed the study) was focused on during re-scoring. To calculate reliability, each individually scored probe was calculated, and then all of the probes were combined to find the average. The experimenter randomly selected 33% of the talk aloud transcriptions and reliability came to 92.5%.

Social Validity

The experimenter provided teachers with copies of questions for each student, as well as a sheet with teacher questions. Students were asked by their classroom teacher after the completion of the study to avoid pressure to answer the questions positively in front of the experimenter. The questions were as follows: First, do you feel the instruction helped you in school? How? Second, Did you like the time during school you were pulled out of your day for instruction in talking aloud? Third, do you think this instruction could help other students your age? Why? Fourth, What do you think could be added that would help you more? And the last question, Would you want to work on this skill again?

The teacher questions were as follows: Teacher questions First, do you feel like this intervention is beneficial for students with disabilities? In what area? Have you noticed a difference in any other content? Third, should we have chosen this behavior as a goal for the
intervention? And the last question asked to teachers was, if the probes and intervention were already made, could you see yourself adding this to your classroom instruction if time was made available?

**Experimental Procedures**

**Baseline.** Students began baseline after parental consent and initial testing. The first student who entered the intervention was determined by the experimenter. The students’ behavior in baseline was measured for at least five data points (What Works Clearinghouse, 2011) using the dependent variable (i.e., same directions as during the intervention). Similar to a previous study (Witcoski, 2012), the student who began the intervention first had the most substantial deterioration, or evidence of no growth, of accuracy over time. Once the first student reached the criterion for the Accuracy Building Stage, the next student with a stable set of learning data began the intervention. The second through last student entered the intervention when the student before them moved onto the second stage of the intervention.

**Accuracy building.** The first stage of the TAPS intervention was the accuracy building stage, a 10-15 minute lesson on talking aloud the logical order of sentence as shown in Appendix B. Following the scripted lesson, a two-minute timed talking aloud the correct sequence of sentences prompt and then the dependent variable, writing the correct sequence of sentences was completed. The two-minute timed talking aloud the correct sequence of sentences prompt consisted of one, four sentence prompt similar to the dependent variable. To determine student success in the lessons, the two-minute timed talking aloud the correct sequence of sentences was recorded and categorized.
The three categories of talk alouds were rereading, talk alouds about order, and talk alouds about cue. Rereading was point-to-point correspondence to the text. Students were able to score one point for each of the sentences (e.g., The fox ran down the road) or 1/2 point for part of the sentence (e.g., the fox). Talk alouds about order were any independent clause about order without any explanation (e.g., this sentence goes first). In other words, the talk alouds about order only stated a position the sentence was arranged in. Talk alouds about cue were any independent clause that used part of the sentence and words to give explanations for decisions (e.g., this sentence explains the other sentences). The talk alouds about cues were anything that lead the student to get to an answer including statements of misunderstandings or unknown information (e.g., I do not know what this word means, I am not sure where this sentence goes so I should keep reading).

When a talk aloud about cues was combined with a talk aloud about order, it was counted as two talk alouds about cues (e.g., student says “The sentence I just read will go last because it explains the other sentences”). A talk aloud was counted as incorrect if it was unrelated to the content in the sentences or unrelated to problem solving. During the Accuracy building phase, no feedback was given following the two-minute timed talk aloud the correct sequence of sentences prompt. The Accuracy building stage ended when students reached half of the exit criterion of the study (i.e., eight talk alouds about cues). The goal of the Accuracy building phase was to instruct students on talking aloud about their decision making and increase explanations.

**Frequency building.** The second stage of the intervention began when students were able to practice the skills learned correctly in the accuracy building stage. To enter the second
stage of the intervention the student was required to talk aloud about cues eight times in two minutes (i.e., half of the exit criterion of the study). The goal of the frequency building stage was to practice problem solving with feedback. Frequency building involved repeatedly practicing a timed behavior followed by immediate performance feedback after the practice trial (Kubina & Yurich, 2012). The student continued practicing over time until he or she reached a criterion or quantitative marker predetermined by the experimenter as developed in prior studies.

Students began frequency building with the task of talk aloud the correct sequence of sentences for two minutes with one minute of feedback and repeated the task again. Feedback began with praise recognizing something the student added to their talk alouds previously not observed or something correctly (e.g., recognizing that they need to re-read a sentence before making a decision). Feedback on incorrects, missing information, or partial answers was given for the remainder of the minute by the experimenter through modeling or asking the student to produce the correct response. The dependent variable was administered following the second one-minute of feedback. The students exited the frequency building stage once they reached 16 talk alouds about cues, or the frequency aim. The frequency aim was established by collected talk aloud data from two “expert” learners in science (i.e., two high school students enrolled in Advanced Placement or AP courses in science) and implemented in a previous study (Witcoski, 2012).

**Maintenance.** Data were then collected once every two weeks for one month following the first week check. The experimenter administered the dependent variable for two minutes. No feedback was given in maintenance because the student independently solved the dependent
variable. The extent of feedback was praise for completing the dependent variable. After completion the student returned to the previous classroom setting.

**Generalization.** Data to determine the generalization of the skills learned in this study using science text were collected throughout the baseline and intervention phase. The experimenter developed probes similar to the dependent variable using Social Studies text (i.e., also text used in a fourth grade classroom). Every five days of instruction or baseline probes, the students were given the generalization probe. Five days was chosen because it represented one school week of instruction.
Chapter 3

Results

Results from the experiment are presented in graphical format. The experimenter displayed data on five tiers taken from the Standard Celeration Chart (Pennypacker, Gutierrez, Lindsley, 2003; White, 1986). One advantage of the SCC-derived figure was to illustrate learning that occurred in real time accounting for all days the student was present and absent due to illness, weekends, or holidays. The figures also display successive calendar days on the horizontal axis and count per minute frequencies on the vertical access with a ratio scale. In other words, each distance on the vertical access of the chart shows equal ratios of change (e.g., distance from 1 to 2 same distance as 10 to 20, both show a doubling of change). With equal ratios of change, the SCC-derived figure presents observable changes as proportional (Graf & Lindsley, 2002) to allow for visual analysis across phases to detect changes in trend and level (Kennedy, 2005). The SCC-derived figure also facilitates the quantification of behavior and behavior change with measures such as celeration (e.g., behavior increased x2.0 over 20 days) and frequency jumps (e.g., upon entering instruction the behavior jumped up x3.0 from baseline to the first day of instruction).

Using celeration, the significance of behavior change is demonstrated by daily performance frequencies changing over time, or a unit of measurement showing count per time unit over a time unit (Johnston & Pennypacker, 2009, Pennypacker et al., 2003). For example, a x1.0 celeration means that the student’s behavior has remained consistent or the trend is flat and not changing. If the goal is for improvement, or acceleration, something needs altered such as
the implementation of an intervention. The experimenter calculated a trend line by using the quarter-intersect technique (Pennypacker, Gutierrez, & Lindsley, 2003).

Another SCC measure used to analyze behavior change is the frequency multiplier (i.e., also known as frequency jumps), or the amount of change between two frequencies. Single case design calls for examining the immediacy of change when an intervention is applied (Cooper, Heron, & Heward, 2007; Kratochwill, Hitchcock, Horner, Levin, Odom, Rindskopf & Shadish, 2013). A frequency jump value (i.e., jump up, jump down, no jump) is found by measuring the distance from the first frequency data point to the second frequency data point (Pennypacker, Gutierrez, & Lindsley, 2003) and quantifies the impact of change. The corresponding value results from the multiply distance on the scale. As an example, going from 1 to 2 would represent a x2.0 jump up, or a doubling of performance.

The last measure used to aid in data analysis was the Accuracy Improvement Measure (AIM) - a measure that accounts for the changes in the quality of a behavior (Pennypacker et al., 2002). AIM is a ratio derived from concurrent correct and incorrect celeration lines. AIM quantifies accuracy or quality of behavior and has published significance criteria (Kubina & Yurich, 2012). The value of AIM demonstrated how much the student’s behavior changed: more or less accurate for the measured data set. All charts used standardized conventions of dots for correct responses and X’s for incorrect responses. Figure 1 displays all student data for baseline, intervention (TAPS and frequency building) and maintenance phases with a counting time of two minutes. Figure 2 shows all of the generalization probes across students in baseline and intervention phases.
Correct and Incorrect Sequence of Sentences

**Abbie.** Figure 1 displays data from Abbie’s results across baseline, intervention (i.e., TAPS and Frequency building), and maintenance phases for sequencing sentences. During baseline, five data points, Abbie showed a consistent worsening for correct sequence of sentences or CSS with a celeration of $\div 2.25$ [10 days] and incorrect sequence of sentences or ISS decelerating slightly with a $\div 1.1$ [10 days]. CSS was moderately variable in baseline with ISS demonstrating more stability. AIM indicated a significant low quality behavior with a value of $\div 2.05$ [10 days].

Abbie began TAPS and frequency building first out of the participants and displayed an overall change in direction for her correct and incorrect celerations. Her performance in CSS accelerated to a consistent $\times 1.25$ [47 days] with ISS decelerating at $\div 1.4$ [47 days]. Her performance resulted in an AIM of $\times 1.75$ [47 days]. Her frequency jump for corrects was $\times 2.0$ and incorrects demonstrated an immediate jump down of $\div 1.7$ from baseline to TAPS instruction. Abbie demonstrated a stable performance of both CSS and ISS during the intervention. Abbie’s performance continued to maintain over five weeks after TAPS until frequency building was completed. Following the intervention phase, her performance was measured for maintenance. Abbie demonstrated levels similar to the intervention phase with incorrects below corrects (i.e., three weeks and five weeks after the intervention demonstrating four correct and zero incorrect sequence of sentences) and her last two weeks with 100% accuracy.

**Jada.** Figure 1 displays data from the second student who started the intervention, Jada. Her baseline performance indicated a steady decline in performance with CSS maintaining at $\times 1.0$ [17 days] and ISS decelerating as well at $\div 1.5$ [17 days]. Her corrects and incorrects were
both moderately variable in the beginning of baseline. After a week of data collection, the CSS became stable and the incorrects remained moderately variable. Overall, her performance was slowing with AIM demonstrating an improving quality of performance at x1.5 [17 days]. Despite the improvement in AIM calculated by the quarter intersect method, the errors were still all above corrects in Jada’s baseline. Persistent, high level incorrects above corrects prompted the experimenter to select Jada to begin the intervention next.

TAPS and frequency building yielded an improving performance in corrects and incorrects. Jada’s corrects maintained and incorrects dropped with celeration values, respectively, of x1.1 [34 days] and ÷1.55 [34 days]. Jada showed a drop in overall level for ISS as well as a slight increase in level for CSS. Upon entering the TAPS intervention Jada demonstrated a consistent performance of x1.0 CSS with a slight jump down of ÷1.25 in ISS. Her AIM improved from baseline to x1.7 [34 days]. During the intervention CSS and ISS performance was variable. Data indicated a slight improvement in accuracy from baseline, with maintenance data that indicated Jada improved her ability to determine the correct sequence of sentences. She scored zero correct and incorrect one week following the intervention, but then returned to her highest performance during intervention at three and five weeks after instruction, scoring four correct and zero incorrect sequence of sentences both days.

**Saddie.** Once Jada entered the frequency building stage of the TAPS, Saddie began the intervention. Saddie showed a consistent decline in performance with the CSS decelerating at ÷2.15 [28 days] and ISS maintaining at a slightly declining celeration of ÷1.3 [28 days]. The resulting AIM during baseline indicated a deterioration in quality or accuracy at ÷1.65 [28 days].
Her performance during baseline for CSS and ISS was variable. Saddie began the intervention as soon as Jada entered the frequency building stage of the intervention.

Saddie started the TAPS and frequency building intervention and continued for 42 calendar days as shown on figure one. During the intervention her performance improved with CSS staying consistent at x1.0 [42 days] and ISS decelerating at ÷1.35 [42 days]. Overall, AIM showed improvement in accuracy at x1.35 [42 days]. Saddie’s correct sequence of sentences had a jump up of x2.0 from baseline phase once beginning the first day of instruction in TAPS. She also had a jump down of ÷2.5 in ISS upon entering the intervention phase. Although Saddie demonstrated an improvement, CSS and ISS data were variable. The frequency multipliers showed immediate and positive change in the performance of sequencing sentences. A week after the intervention phase, Saddie scored two CSS and two ISS and three weeks after scored four CSS and zero ISS, consistent with her trend in intervention. The data following the intervention demonstrates the effects from the intervention maintained for up to three weeks following the termination of instruction.

**Sasha.** Sasha was in baseline for 44 days with a consistent CSS celeration of x1.0 [44 days] and in ISS accelerating at x1.05 [44 days]. Sasha answered incorrects more often than corrects by an average ratio of 2 to 1. Sasha’s AIM demonstrated a slightly worsening baseline at ÷1.05 [44 days]. Her data were variable for both ISS and CSS in baseline. Sasha began TAPS when Saddie reached the criteria to enter the frequency building portion of the intervention.

During TAPS and frequency building, Sasha’s correct and incorrect celerations reversed. Her CSS accelerated at x1.2 [48 days] and her ISS decelerated at ÷ 1.2 [48 days]. Sasha had a small accuracy improvement during TAPS but overall eliminated most incorrects during the time
in TAPS and frequency building with an AIM of x1.44 [49 days]. Although there is a clear trend in the CSS and ISS data, the data were variable. Her performance on the first day of the intervention slowed with a jump down in both CSS (÷ 1.7) and ISS (÷ 2.0). One data point was collected a week following instruction demonstrating four CSS and zero ISS in maintenance.

**Theodore.** Theodore was in baseline for the longest period of time, 52 calendar days. During baseline his corrects and incorrects maintained. His CSS accelerated at x1.15 [52 days] and ISS accelerated at x1.05 [52 days]. Although both were accelerating, his incorrects were still well above his corrects with a slowly improving AIM of x1.1 [52 days]. Baseline ISS data were stable, with moderately variable CSS data. Once Sasha reached the criterion for frequency building, Theodore began the intervention.

Theodore began the TAPS accuracy building stage with an immediate jump up in his CSS performance of x2.0 correct performance. His ISS also had a jump down of ÷1.3. During the TAPS and frequency building intervention, but with incorrects decelerating more rapidly leading to an increase of quality with an AIM of x1.25 [48 days]. Theodore had an increase in CSS accelerating at x1.25 [48 days], and a steady performance in ISS of x1.0 [48 days]. Although a slight improvement was calculated, both ISS and CSS were variable. Overall corrects increased and incorrects dropped and remained steady resulting in a more accurate performance. The school year ended and no maintenance data could be collected for Theodore.

**Generalization Measures**

Social studies probes were given to the students every five instructional days for two minutes to test for generalization of the skill to another content area. Data for generalization measures are displayed on Figure 2. Abbie had one data point before intervention demonstrating
a frequency jump of x1.3 for CSS and an immediate change with a jump down in ISS of ÷10.0. Overall her CSS maintained at x1.0 [47 days] and ISS at ÷1.0 [47 days] following the intervention. Jada was similar to Abbie in that she only had two data points in baseline. She had a large drop in frequency for incorrects with a jump down of ÷6.0 but her CSS also jumped down ÷4.0. Despite corrects decreasing upon entering the intervention phase CSS continued to increase during the instructional phase at a celeration of x1.75 [34 days] and incorrects at ÷1.15 [34 days]. AIM for Jada was a clear improvement in accuracy of performance across time of x2.0 [34 days].

Saddie was the third student to begin the intervention and she also had few data points in baseline. Her data demonstrated a jump down in both corrects ÷6.0 and incorrects ÷1.25, with celeration of x1.4 [42 days] for corrects and ÷1.25 [42 days] for incorrects (i.e. similar to Jada). Saddie’s AIM on the generalization measure accelerated at x1.75 [42 days]. Sasha had four data points in baseline, but had five in intervention phase. Initially after starting the intervention Sasha’s CSS jumped down ÷3.0, but her incorrects also jumped down ÷2.3 (i.e., similar to both Jada and Saddie). After the intervention began she had a celeration in CSS of x1.1 [48 days] and a deceleration in ISS of ÷1.15 [48 days]. Her celeration measures demonstrated an improvement in accuracy using AIM of x 1.3 [48 days]. Theodore was the last student to enter the intervention. He showed consistency when entering the intervention for CSS with a frequency jump of x1.0 and a jump down of ÷3.0 in ISS. Before starting the intervention his ISS were maintaining at x1.0 [52 days] and CSS were increasing prior to the intervention at x1.25 [52 days]. Theodore’s corrects decelerated at ÷ 1.3 [48 days] and incorrects remained constant x 1.0 [48 days] demonstrating a decline in accuracy overtime of ÷ 1.3 [48 days]
Social Validity

Following the collection of the maintenance data, or intervention data (i.e., for Theodore), participants were asked questions concerning the instruction, time of day they were pulled out of the classroom, outcomes, and future use of the intervention. The fifth participants files were lost by the teacher at the end of the school year. Participants viewed the instruction in talking aloud and solving problems as a relevant skill. Three of the four students stated that the intervention made them better at reading with help understanding bigger words, and what they were reading in the problems. Abbie stated that the intervention helped her take her time and really think aloud about the order. The students did not object to being pulled out of the classroom for instruction and three of the four stated it had helped them with a certain skill (answer questions better, reading new and different things, and realize mistakes). All four students thought other students would benefit from the intervention (having trouble ordering events, trouble reading, or not understanding word meaning). When asked what they would add to the instruction three students said nothing, with Sasha wanting to read aloud more. All four students would want to work on the intervention of TAPS again.

Classroom special education teachers were also provided with a set of questions. Both teachers felt it was beneficial with one teacher stating that the intervention helped students who struggle with communicating what they are thinking. She felt that sometimes they need to be given a script and taught how to “think” out loud and given permission to express themselves in this way. Both teachers did not get much of an opportunity to notice a change in other content because the intervention was implemented toward the end of the school year. Teachers thought talking aloud about problem solving was a relevant goal with one teacher stating verbalizing
thinking was a critical life skill (e.g., what students need help with, and what students don’t understand so teachers can better serve their needs).
Chapter 4

Discussion

The current educational environment in science is not conducive to many students with learning disability’s needs; students with disabilities spend at least 80% of their day in general education (Brigham, Scruggs, & Mastropieri, 2011). The time spent in the classroom, should reflect practices aimed at helping all students reach a level of proficiency with important concepts and skills. Researchers propose instruction should be strongly linked to the content in curriculum in order to improve students’ problem solving abilities (National Research Council, 2000; Pintrich, 2002; Schraw, 1998). Response to Intervention (RTI), a system to respond to inadequate student progress, requires constant monitoring of student performance and research based instructional methods. Schools must now show student gains through instruction, and adjust instruction based on results. Talk Aloud Problem Solving (TAPS) allows for science teachers to monitor student learning by asking all students to talk aloud what they are thinking about science content. The problems should be based within classroom curriculum to help teachers make directed and timely instructional decisions.

TAPS has been shown to benefit problem solving instruction for students in secondary and college settings (Jeon, Huffman, & Noh, 2005; Johnson & Chung, 1999; Pate, Wardlow & Johnson, 2004, Pestel, 1993). The combination of TAPS with frequency instruction demonstrated promise in a study with students at-risk, improving their problem solving accuracy (i.e., sequencing sentences in the correct order) for all participants (Witcoski, 2012). The positive effects of TAPS have led to the current research study that sought to answer the following questions: Does TAPS package of instruction have an effect on a problem solving skill
involving science text? Do students maintain the ability to solve problems overtime following TAPS and frequency building? Does a problem solving skill targeting science content generalize to other content areas, specifically social studies? Do teachers and students view the intervention of TAPS and frequency building as a socially valid and beneficial classroom tool?

The effect of TAPS and frequency building on a problem solving skill was measured using a multiple baseline design across participants. All students increased in accuracy of performance from baseline. Prior to entering TAPS, students demonstrated an inaccurate performance with the level of incorrects above the level of corrects for the length of baseline. Four of the five participants were becoming less accurate in baseline, some significantly deteriorating such as Saddie with an AIM of ÷1.65 [28 days] and Abbie, AIM ÷2.05 [10 days]. The other student, Theodore had slightly improving accuracy, but the level of his incorrects stayed higher than his corrects throughout baseline (i.e., 52 calendar days) meaning he had consistently more incorrect answers to correct when solving problem.

The degree of change in AIM shows remarkable consistency with all students improving the quality of their answers ranging from x1.25 to x1.75 better. The quality improvement measured by AIM indicates every student positively changed their ability to sequence sentences with mixed science content. The magnitude of significance change follows: Theodore (x1.25) small accuracy improvement, Saddie (x1.35) and Sasha (x.145) adequate accuracy improvement, Jada (x.17) and Abbie (x1.75) substantial accuracy improvements. Additionally, for four out of five students the baseline for accuracy demonstrated that even with general education science instruction, the students did not improve through classroom instruction or maturation. The TAPS and frequency building intervention had a positive effect on the quality of problem solving. The
data reflect similar findings by Pestel (1993) showing that students introduced to TAPS answered fewer problems completely right, with the other group getting more problems completely right or wrong. Early stage TAPS learners appear to have a tendency to try to work through the problem rather than either knowing or not knowing the answer. By the end of the study students were demonstrating a promising understanding of textual features to help aid in organizing paragraphs.

TAPS and frequency building was introduced to all students, with only four reaching the exit criterion (16 talk alouds about cues). Frequency multipliers were used to quantitatively examine the immediate change in student performance once instruction for the intervention began. Upon introduction to the intervention, Saddie, Abbie, and Theodore all had substantial jump ups in corrects with a doubling in performance (i.e., x2.0). The x2.0 frequency multiplier for corrects indicates the intervention had a strong immediate impact in students ability to problem solve. The other two students had no jump (Jada, x1.0) and a jump down (Sasha, ÷1.7). However, Jada incorrects jumped down by a factor of ÷1.25 and Sasha’s incorrects jumped down by ÷2.0. When placed in context, the TAPS and frequency building intervention produced still a positive effect for for Jada and Sasha.

The study demonstrated improvement was consistent overtime through the celeration measures. The celeration for all students prior to the intervention showed the magnitude of need to begin instruction. Abbie and Saddie both had significant reductions in correct answers with, respectively, ÷2.25[10 days] and ÷2.15 [28 days] decelerations. While incorrects were also decelerating, Abbie ÷1.1 [10 days] and Saddie ÷1.3 [28 days], the corrects outpaced than the incorrects in loss of learning. Jada had decelerating incorrects in baseline, however, only one day out of nine had incorrects at the same level of corrects with the remainder of incorrects at a
level above corrects. Sasha and Theodore both demonstrated that without instruction, the long
trends of near flat lines, or no change in inaccurate performance, would continue for a large
portion of the school year. The significance for Sasha's corrects, x1.0 and Theodore, x1.15 both
fall in the range of unacceptable growth celeration (Kubina & Yurich, 2012).

Concurrent celerations also indicated that during the intervention all students’s learning
were improving. Abbie, Jada, and Sasha had accelerating corrects and decelerating incorrects
creating a clear picture of improvement. Saddie had corrects and Theodore had incorrects that
remained flat throughout the intervention; despite no change, the other measure incorrects (i.e.,
decelerated) or corrects (i.e., accelerated) both improved. The celerations demonstrated the
robustness of the intervention and how it differentially affected each individual.

Previous TAPS research demonstrated students referred to concepts taught more often
than students not using TAPS (Glass, 1992), correctly completed more problems (Jeon, Huffman,
& Noh, 2005; Johnson & Chung, 1999; Pate, Wardlow, & Johnson, 2004) and correctly
evaluated more of their hypothesis and had fewer incorrect evaluations when troubleshooting
electrical systems with faults (Johnson & Chung, 1999). The present study also showed benefits
in TAPS over time. All student performances during the intervention stayed consistent or
remained better than in baseline.

With a multiple baseline design, experimental control is determined when an intervention
is applied to a tier and no spillover of effects occurs in subsequent tiers (Kennedy, 2005). The
intervention must also show significant effects for changes in trend, immediacy of impact, and
consistency in data pattern with subsequent phases (Kratochwill et al., 2013). The present study
shows a positive immediate impact for all students when introduced to the TAPS and frequency
building intervention. Furthermore, the celeration for correct and incorrect sequences of science content improved for all five students. And last, the Accuracy Improvement Measure (AIM) demonstrated a positive change for all five students. The changes from baseline produced by the TAPS and frequency building intervention indicate a functional relation.

Maintenance data were collected for four of the five students. Jada was unable to answer any of the problems during the first maintenance check (i.e., zero correct and incorrect). The experimenter asked her to give directions following this performance in order to ensure she was aware of the task (i.e., she immediately jumped up to four correct and zero incorrect for the remainder of data points). Abbie, Saddie, Jada, and Sasha all maintained one of their best performances (i.e., four correct, zero incorrect) following the TAPS intervention and frequency building. Theodore did not finish the intervention by the end of the school year and no maintenance data were collected. Overall, the maintenance data demonstrated that the students applied the skills learned during the intervention even after the instruction stopped. The maintenance data support the proposition of behavioral fluency that once a behavior reaches a frequency standard (also called a fluency aim) behavior persists for long periods of time after practice has terminated (Binder, 1996; Kubina, 2005; Kubina, Amato, Schwilk, & Therrien, 2008; Kubina & Yurich, 2012).

Along with maintenance, all students completed generalization probes in social studies content during baseline and the intervention phase. Abbie demonstrated a jump up and a sizable jump down and maintained her performance across time with $x1.0 \ [47 \text{ days}]$. Theodore, like Abbie, also had an improvement immediately upon entering the intervention but then started to decline in accuracy or AIM over the course of the intervention phase at $\div 1.3 \ [48 \text{ days}]$. The
other three students had similar results for generalization measures. Jada, Saddie, and Sasha had immediate drops in both corrects and incorrects upon entering the intervention. The three girls also had improvements in accuracy overtime with Jada at x2.0 [34 days], Saddie at x1.75 [42 days] and Sasha with x1.3 [48 days]. Although there were a lack of data points, visual analysis indicates a change in level of corrects and incorrects (i.e., corrects at a higher level and incorrects at a lower level than in baseline). Overall the intervention had a positive effect on generalization measures for four of the five students. Future studies should include more opportunities to collect data, specifically in baseline, to determine a more conclusive effect on other content areas.

TAPS with frequency building was shown to have a promising effect on problem solving when applied to science content. The intervention was also seen as beneficial by students who enjoyed the intervention and teachers who thought it was worthwhile. The study ran for over 100 days, allowing enough time for students to form an opinion of the frequently used problem solving method. Despite the length of time, students liked using the techniques that they learned and would want to continue its use in the future.

**Implications**

TAPPS and frequency building demonstrate a way for teachers to teach a thinking skill. Namely, understanding textual features and organizing paragraphs along with language skills and solving problem in science. Not only did it benefit the students when comprehending expository text, but also provided students with immediate and relevant feedback to communicate more effectively. The skill of communication in problem solving is essential to engage in the practices and language of science (e.g., analyze data, construct explanations) and to form a deeper
understanding of science (National Research Council, 2005). The current measurement of daily instruction in varied curriculum is poor in providing guidance for teachers (Stern & Ahlgren, 2002), making the ability to monitor student mastery more difficult.

A measurement technique that is sensitive enough to detect changes in daily instruction is frequency, or accuracy and speed of a behavior (Howell & Lorson-Howell, 1990). When a behavior becomes fluent it can be described by words such as effortless, easy, accurate or automatic. Experts in a behavior such as problem solving are easily detectable by observers when compared to novices. In the current study, the performance of high school advanced placement science students, was used to determine a possible fluent performance determining expository text structures and sequencing sentences into a logical order (speed and accuracy measures). Frequency building to a performance criterion was established using “expert” learners, and applied to instruction for younger students. The experimenter used behavioral fluency (Binder, 1996; Kubina & Yurich, 2012) as an outcome that can be achieved through explicit practice procedures known as frequency building. A performance standard is a vital benchmark, with a large body of research showing certain associated critical learning outcomes occur: long-term retention, endurance or resistance to fatigue, and application or the ability to apply element skill(s) to a more complex compound skill (Binder, 1996; Kubina & Yurich, 2012).

In a previous study, increasing the frequency of statements about cues and explanations of sentences increased the accuracy of solving problems (i.e., sequencing sentences in the correct order) for students labeled at-risk for learning disabilities (Witcoski, 2012). Previous studies in reading (Kubina et al., 2008; Brown, Dunne & Cooper, 1996), mathematics (Binder, Haughton,
Van Eyk, 1990; Brady & Kubina, 2010) and writing instruction (Datchuk, 2011) have successfully used behavioral fluency to a FBPC. Overall results from the current study, demonstrate that the intervention using behavioral fluency produced a change in problem solving behavior that endured and maintained across time. For the majority of students, the effect also generalized across content areas to social studies.

**Future Directions and Limitations**

Talk aloud problem solving and frequency building holds promise in content area instruction. Additional exploration of the TAPS intervention is needed to determine more precise and efficient ways to allow for use in the classroom. The current study replicates Witcoski (2012), but additional data are needed to examine how the intervention can be used with a variety of populations and students. All students were required to meet the exit criterion (16 talk alouds about cues) prior to leaving the intervention phase, but in future replications more exploration should be focused on evaluating varied performance standards. Because performance standards are well researched criteria that indicate a fluent performance, fluent levels are under scrutiny until further research is available.

The current study also displayed high variability in some data. Because of the nature of placing sentences in logical order, the answers directly affect one another. In other words, if a student gets the first sentence incorrect, they will automatically get another sentence incorrect that should be in that order (e.g., students says the order is 1,3,4,2 but the order should be 2, 3, 4,1). Future studies should explore different ways to help lessen variability for specific problem solving types and increase sensitivity. Along with variability of data, transfer of talking aloud to
writing answers should be addressed in future studies (i.e., seeing the prompts and saying the answer compared to seeing the prompts and writing the answer).

The current study explored TAPS and frequency building for students with disabilities and is the first to explore the use of the TAPS package of instruction for students diagnosed with a disability. All of the students in the intervention experienced a drop in performance during the intervention that should be evaluated more in future research studies. Future studies should also have generalization probes implemented more frequently, to allow for a more sensitive measure. Future studies might also explore the intervention with student pairs (i.e., one student as listener and the other as problem solver) rather than only the teacher serving in the role of expert. The study adds to the current literature working with older students in more advanced science course, a girl with autism, and students at-risk for disabilities. Replications are necessary to determine the effect on different ages, ability levels, and settings. Along with replications, more fielded tested probes would be beneficial for a more random selection of instruction materials.
References


University of Minnesota). Retrieved from ProQuest Dissertations & Theses database.

(UMI No. 303979616).


Mastropieri, M. A, Berkeley, S., McDuffie, K., Graff, H., Marshak, L., Conners, N., ....


Unpublished manuscript, Department of Special Education, The Pennsylvania State University, University Park, PA.
Baseline TAPS and frequency building
Feb 24 Mar 24 Apr 21 May 19

Count per 2 minutes for Sequences of Sentences
Generalization

Abbie
Jada
Saddie
Sasha
Theodore

Figure 2
Appendix A

Literature Review

Introduction

Students across the country are struggling with complex reading tasks as reflected by national test results. The 2011 National Assessment of Educational Progress (NAEP) reading results for 4th and 8th grade indicate 34% and 25% falling below basic with only 34% and 43% performing at a basic level. By the time students reach 4th grade, they are expected to begin reading more information in the form of increasingly difficult expository text in content areas. The transition is complicated by the fact that students with disabilities have profound issues reading expository text because of more complex text structure, level of complex vocabulary, conceptual density, and importance of prior knowledge (Saenz & Fuchs, 2002). Reading is already a difficult task that requires a grand synthesis of many skills and strategies (Carnine, Silbert, Kame’enui & Tarver, 2009). Overtime, students who are struggling will fall further behind as content areas become more complex and rely more heavily on textbooks for learning.

Mathematics is important to recognize in science because of an increase in quantification, visualization, and precision (Kline, 1980). Therefore, reading and progressively more mathematics are involved as a student reaches advanced sciences. Results are also concerning for mathematics with 18% and 28% in 4th and 8th grades falling below basic. A larger portion of students functioned at the basic level in mathematics with 42% and 39% of 4th and 8th grade students. The combination of reading and mathematical skills (i.e., increasing quantification of science) necessary to become proficient is reflected in national Science results. By 8th grade,
35% of students do not meet the basic level in science, with only 33% performing at a basic level (NAEP, 2011). The data for reading and math indicates an area of intense need for rigorous study of problem solving in science and other difficult text and content. The portion of the student population that struggles the most is special education.

A strong understanding of science is a goal for every student, including those with special needs. Science encompasses natural systems and processes with knowledge that is established and continually changing and growing. Science is composed of several different levels of understanding of an ongoing process, from the facts that are integrated and articulated into well-tested theories, to bodies of data and outcomes of experiments explained by theories, along with the limitations of those theories (Duschl, Schweingruber, & Shouse, 2007). Much of what is currently being taught in science classrooms is not compatible with the instructional needs of students with special needs, specifically those with learning disabilities (Brigham, Scruggs, & Mastropieri, 2011). Considering that students with disabilities spend at least 80% of their day in general education, more attention is needed on what type of instruction is occurring in these settings. Based on what traditional instruction focuses on, students are not receiving the instruction that is necessary to promote understanding and comprehension of complex expository text (Mason & Hedin, 2011).

The statistics of student performance are particularly disturbing because in advanced science courses, science is taught primarily through lectures and textbooks (Brigham, Scruggs, & Mastropieri, 2011) that are dominated by facts and algorithmic processing (DeHaan, 2009). The traditional science instruction approach of relying on lecture and textbooks is not sufficient in promoting student understanding of material not yet mastered or introduced. Students are
expected to comprehend the text and develop their own representations from the information.

Students must also gain understanding that they need to add to their current knowledge, construct new understanding, confirm what they already know, restructure knowledge and other stages that come with new experiences and instruction. During the instructional process, students with disabilities may struggle during the early stage of learning with the text (e.g., complex vocabulary or lack of prior knowledge) among other things. The inability to instruct a variety of students leads to outcomes such as disengagement, with students with disabilities rarely entering the STEM workforce, though many are capable of making valuable contributions (Leddy, 2010).

Evidence is available that with appropriate instructional support and structured experiences, students from varied socioeconomic, cultural and linguistic backgrounds are able to engage in science, generating and evaluating scientific evidence and explanations (Kuhn & Dean, 2005; Lehrer & Schauble Metz, 2004; Roseberry & Warren, 1994). The improvements are not simply in basic concepts; researchers found when given the opportunity to engage in more sophisticated scientific experiences such as investigations (Lehrer & Schauble, 2005; Metz, 2004), exposure to appropriate experiences is vital. All parts of science can be taught to students included in general education content.

One theme remains consistent across research - appropriate instruction and structured learning experiences are vital for student success. Students with access to previously mastered content, experience in scientific processes and dialogue, and quality instruction and feedback will have an advantage when exposed to more difficult content. Within subject areas (e.g., history, science) students are expected to acquire and integrate thinking skills in authentic problem
solving activities (Kameenui & Carnine, 1998). The expectation of success demonstrates that educators need to do more to prepare students to fully engage in the classroom.

Current literature demonstrates a lack of research for tools teachers can use to measure student progress more frequently to help guide classroom instruction. Specifically, frequency measures would be beneficial as a tool to discern students who have mastered concepts and skill from those who are struggling. Along with frequency measures, the benchmarks determining what fluent performances look like would assist teachers. Stated differently, how frequency measures lead to the identification and attainment of fluency is valuable for teacher and students. Fluency or automaticity increases the likelihood that a student has critical learning outcomes: long-term retention, endurance, application (Binder, 1996; Kubina & Yurich, 2012); outcomes all necessary for a deep understanding of science.

**Problem Solving**

Problem solving is involved in science and every other content area in school and daily situations in society. The term has been functionally defined as “any behavior, which through manipulation of variables, makes the appearance of a solution more probable” (Skinner, 1953, p. 247). Skinner’s definition is inclusive to all content areas and approaches problem solving as an active and ongoing process. Problem solving instruction begins in the classroom, making it imperative that teachers allow for ongoing feedback and practice actively engaged with problems.

Problem solving can be seen in classrooms through activities such as strategy use, self-monitoring, visual imaging, and engaging in covert behavior such as thinking (Palmer, 1991). One method that holds promise using a combination of typical classroom behaviors for developing skills used in problem solving is TAPPS or Talk Aloud Pair Problem Solving (i.e.,
also known as TAPS). The learner engages in problem solving while explaining every step in reasoning and the listener has to understand and use questioning techniques for every step the problem solver takes (Whimbey & Lochhead, 1999).

Talk aloud problem solving is a method to teach students the language of problem solving while continually working on skills necessary for academic success (Whimbey & Lochhead, 1999). Through talking aloud students are encouraged to keep a positive attitude about skills they have not yet mastered. Problem solving requires students to attempt different methods to find a solution that may require re-reading, adjusting procedures, and understanding that some complex problems require more time to solve. Accuracy is a main goal of TAPS, requiring a student to develop a full understanding of the problem and rechecking that all parts are correct (1999).

Problem solving can range from a simple two step problem to a complex ten step formula. To fully understand the components of a problem, each part of the problem needs to be pulled apart and placed into simpler steps. By checking for accuracy and breaking a problem into the smallest parts, students avoid guessing and attempt to come to a conclusion. Good problem solvers understand that activeness during problem solving and talking the problem aloud with oneself and others is helpful for a better understanding and accuracy (Whimbey & Lochhead, 1999). With all of the skills necessary to become a proficient problem solver, having students talk aloud their thinking is a beneficial and straightforward method for determining what will contribute to better problem solving for each individual student.

Talk aloud problem solving (TAPS) also known by Talk Aloud Pair Problem Solving (TAPPS) has been developed over the years to include some guides to create activities to help
students. One learning goal used to measure student use of strategies during problem solving was through anomalies in reading (Whimbey, 1985). For example, text structure functions as an important cueing system to communicate logical connections and relation of ideas to others (Meyer, Brandt, & Bluth, 1980). Changing the structure of the text creates an exercise which triggers the finding of errors, while reading triggers the ability to learn from text and use appropriate learned strategies to solve the problem (e.g., using science text and taking the process of the water cycle out of order, students need to use knowledge of transition words, prior knowledge of the water cycle and noun phrases to guide their answer). The knowledge needed to understand text structure makes it important as an underlying comprehension skill (NAEP, 2011). Before deciding which strategy will fit the purpose of the text or problem, students must be aware that their understanding of the text is inaccurate or inadequate. Readers initiate repair of comprehension when they “detect inconsistencies between their knowledge and external information…or between elements of external information” (Otero, 2002, p. 283) that is the internal consistency of the text.

Despite the prominence of problem solving, students are not often taught to explicitly problem solve, but are expected to select the correct use of language from lectures or lessons with very little feedback. There is a clear difference in the ability of an expert compared to a novice, and teachers can collect data on successful student performance to help those that struggle. There is evidence to support that students with disabilities are lacking the ability to determine different text structures, affecting their ability to read and write coherent prose (Englert & Thomas, 1987). The complex skill of solving problems is intertwined with skill based problems on exams of student achievement and is a challenging method to learn.
Problem solving is an area of importance across disciplines and viewpoints. The language of problem solving should be used to teach others to solve problems as well as how to provide immediate feedback on a performance. Learning to verbalize problem solving steps can help students practice communication skills and receive corrective feedback on performance. Several theories of how students problem solve have emerged from the literature. Despite these differences, how teachers measure a student's success to problem solve remains the same (i.e., testing, verbal repertoires, daily living skills).

Talk aloud problem solving (TAPS) is aligned with a model of instruction that allows teachers to have students actively involved in science. The activities are overt, allowing for constant monitoring of strategy use and measurement of problem solving abilities (Whimbey & Lochhead, 1999). Teacher behavior could be influenced by student responses to provide direct and explicit feedback along with instruction for appropriate strategy use when needed.

Strategy instruction combined with direct instruction, does improve student performance for students with learning disabilities (Swanson, 2001; Swanson & Hoskyn, 2001). The evidence in the special education field demonstrates that it could be used in the general education classroom more broadly with hard to teach skills. Because problem solving can be broken down into smaller component skills, a direct and explicit design to instruction would seem necessary for students struggling to master the problem solving process. With the absence of a measurement instrument for problem solving abilities, the field is lacking in the tools for teachers to use to guide decision making for student needs. TAPPS is a technique that can provide a model for teachers, and has established popularity within the science field and leaders in the field of education.
The current review was conducted to pursue the answers for three questions: What instructional components did studies use to implement TAPS in the science classroom? What does research include about teaching TAPS to students functioning at different levels of need? What effect did TAPS have on the problem solving performance of students? And what are the current practices in science education?

**TAPS Studies**

Talk Aloud Pair Problem Solving has been implemented in science classrooms for use in teaching problem solving concepts. Although the use of TAPS (i.e. also referred to as TAPPS) is encouraged in the science field (Behham, 2009; Bodner & Herron, 2002; Brent & Felder, 2012; Lee Kam Wah, 1998; Tingle & Good, 1990), studies are lacking. Only six studies could be located using students in older grades, including a dissertation and a case study.

The ages ranged from middle school (i.e., 7th and 8th grades) to college level courses. Three of the six studies took place in a high school setting (Glass, 1992; Jeon, Huffman & Noh, 2005; Pate & Miller, 2011) with one study also taking place in a middle school (Glass, 1992). The remainder of studies occurred in a college setting. Two of the three studies used TAPPS instruction in chemistry content (Jeon et al., 2005; Pestel, 1993), one in aviation technician class (Johnson & Chung, 1999), another in career and technical education courses (Pate & Miller, 2011) and the last two studies used TAPPS in technology education courses (Glass, 1992;; Johnson & Chung, 1999; Pate & Miller, 2011).

**Summary of TAPS Measures and Results**

All of the studies tested the effect of TAPPS on problem solving. Four of the six studies examined the effect of TAPPS on a specific problem solving task. The other two studies
investigated grade distribution and patterns of performance on an exam (Pestel, 1993) and essay scores (Jeon et al., 2005). Despite different measures, all of the studies measured a form of problem solving in science. Three of the studies also examined verbal protocols for additional analysis (Glass, 1992; Johnson & Chung, 1999; Pate & Miller, 2011). The verbal protocols were used to determine what language the students were using and comparing the results against what is known about good problem solvers.

A time component was added to Pate and Miller (2011) for all students and Pate et al. (2004) for only successful students, determining time to task completion. Prior knowledge or pretests were used to determine student equivalence and participation for all studies. All of the studies used video or audio recordings with the exception of Pestel (1993). Audio or video is beneficial for reliability and fidelity measures, however none of the studies specifically outlined what they did to check for fidelity of treatment or scoring reliability. Only one study had students complete a social validity survey (Johnson & Chung, 1999).

**Chemistry studies.** Pestel (1993) studied the effects of TAPPS on chemistry problem solving. The students were two first semester general chemistry courses at the Rose-Hulman Institute of Technology. The case study was not randomized, but included a control group with a conventional modeling strategy and the group using TAPS. The task outcome was determined by looking at problems in terms of completeness (i.e., completely right, partially right, or incorrect). The scoring eliminated issues with some of the students that might have been marked incorrect but still had some of the problems correct, versus students who gave up prior to completing any parts. The TAPS class answered fewer problems completely right, but they also
answered fewer completely wrong. The conventional method showed a greater tendency to get problems completely right or wrong.

The second study that looked at chemistry problem solving was Jeon, Huffman, and Noh (2005). The participants were 11th grade students from a boys school in Seoul randomly assigned to groups. Students either used a problem solving strategy individually (i.e., a 4 step strategy), TAPPS with a script, or the control group with a text-book approach. All groups were taught by the same teacher, who practiced prior to teaching each group. Chemistry and math backgrounds were analyzed and the differences between higher and lower performing student pairs were consistent across pairs. The posttest included transcribed answers to essays that were audio and videotaped. The mean score for all TAPPS groups was higher than all other groups. Both individual and TAPPS performed better than control in the essays, with TAPPS scoring better than all other groups when looking at conceptual knowledge (i.e., the conceptual knowledge did not improve for the two other groups). Overall the students scored 66.7% answering correctly in the TAPPS group compared to 40.0% and 14.3% in the individual and control.

**Technology education studies.** Glass (1992) was the only dissertation to meet the inclusion criteria. The study included three technology education classrooms and laboratories at a junior and senior high school. The study had a TAPPS group, no TAPPS group, and a control group (i.e., no talking aloud) in a randomized block design, all naive to the concepts in the study. A pilot study was conducted to determine the appropriate amount of time needed for problem completion along with data collection procedures and problems. Students were tested with Iowa tests of basic skills with no significant differences found, and were given an instructional unit
and mastery test. Remediation was provided for those who did not meet the mastery requirements for prior knowledge until they succeeded. All four problems created for the dependent variable included construction components from technology education subject matter. Verbal protocols were also analyzed for all students.

A practice day began the study to rehearse using TAPPS with the researcher, but was not counted toward the results of the study. The teacher and researcher were present and gave reinforcement and time warnings for the activity. The researcher explored several factors within the data set from this study including: student understanding of the problem including facts, elements and limitations (i.e., two way ANOVA, F ratio significant at .05 level), amount or degree of transfer of concepts (i.e., means and count of statements), degree of metacognitive strategy or skill use (i.e., content analysis of several separate variables), quality of construction solutions (i.e., four separate two way analyses of variance), and consistent behavior differences in problem solving (i.e., total number of statements and categories). The TAPPS group mentioned more statements along with more concepts learned during class. A significant degree of difference was found in the amount of strategies, developed statements including explanations, comprehension monitoring and evaluation, along with implementing modifications to the solution. Although there was improvement in the verbalization of statements, no difference was found in how well the students understood the elements, facts and limitations of the problem. Overall, the TAPPS group was referencing the problem and ways to solve it, resulting in more focus.

Johnson and Chung (1999) implemented a TAPPS intervention in two college classes teaching second level electronic systems. Twenty-eight students were measured using verbal
protocols and task completion in a non-equivalent control group design. All participants completed prerequisite courses, participated in laboratory exercises, and performed no differently on domain knowledge tests. T tests were completed to determine the differences between the groups on the variables of problem recognition, correct solution, practice time, and simulated work time. TAPPS students were significantly more successful in fault detection, identification of faulty components, and evaluation of faulty hypotheses than students in the control group.

Expanding on the previous study, Pate, Wardlow, and Johnson (2004) ran a study with 30 college students in a first semester small power technology course. The dependent measures were task outcome (i.e., successful or unsuccessful) and time completion on the post-test for successful performances. All students were randomly assigned to control or experimental groups with a counter-balanced internal replication. No differences were found on pretests after receiving instruction and laboratory practice. TAPPS doubled the success rate at solving small engine faults, supporting the results found in Johnson and Chung (1999). Both groups were audio taped with no significant difference in completion time. When reversing roles students did not transfer the strategy which resulted in lower task completion when talking aloud prompt was removed.

Pate and Miller (2011) conducted a recent study in high school career and technician classes with TAPPS with 16 students across four Iowa schools. The study used a randomized post-test only control group design seeking to explore if there was an optimum level of statements conducive to successful troubleshooting. Students in the TAPPS group practiced using the technique twice prior to the study. In total, four students out of 16 using TAPPS were successful, with no support shown for TAPPS. Whereas the control group demonstrated that
seven out of 18 were successful. Students that successfully completed the troubleshooting task showed no significant difference in mean time to complete the task between groups. Unlike Pate et al. (2004) student instruction Pate and Miller (2011) only required one class, compared to a prerequisite course related to the topic (Pate et al., 2004)

Recordings in Pate and Miller (2011) demonstrated students who were unsuccessful at problem solving used twice as many self-assessment statements and not-on-task statements. Besides the self-assessment statements discovery no minimum level of statements were found to lend itself to successful performances because of high variation in performances. The primary discovery was the TAPPS strategy does not improve performance when secondary-level students do not have enough domain specific knowledge.

**TAPS Instructional Components**

Overall, studies demonstrated the importance of several criteria when implementing TAPS in the classroom. Teachers or experimenters should adjust times based on how long students need to complete a problem and try out different room arrangements when working in pairs (Glass, 1992). Types of problems and student population may require extra time based on the amount of practice the participants have received.

The reviewed studies offered several ways to help prepare students to use TAPS in the classroom. The switch to active problem solving (i.e., receiving and providing constant feedback from peers) requires some preparation such as practice with easy problems to learn the format of talking/listening (Whimbey & Lochhead, 1999). Practice with TAPPS was specified in four studies (Glass, 1992; Jeon et al., 2005; Johnson & Chung, 1999; Pate & Miller, 2011; Pestel, 1993). A script can be helpful when learning TAPS procedures (Jeon et al., 2005) and can
facilitate discussion that is aimed at on-task behaviors. Students may not have had experience with the technique and practice with prompts can assist when working with peers.

All but one study (Pestel, 1993) used some type of recording to score student performance following the intervention. Despite Pestel (1993) not using a recording, permanent product was used for the dependent variable, allowing the researcher to re-score if necessary. However, procedural fidelity should be questioned as it was not reported. None of the studies specified what they did to re-score or check for treatment fidelity, causing concern for accurate and reliable scoring.

Pre-testing for prior knowledge without a specific criteria can add additional confounds making TAPS less effective on problem solving (Pate & Miller, 2011; Pestel, 1993). Students that are unaware of basic concepts being tested will not be able to successfully complete the work when placed into pairs. Also, the time provided should be adjusted based on students, with only Glass (1992) solidifying the amount of time needed (i.e., using a pilot study) for the specific task given. This may be an issue in some studies, if students could solve the problems, but not in the time provided.

Effects of TAPS on Problem Solving

More than half the studies using TAPS for problem solving in science demonstrated an effect of TAPS on problems solving measures. An interesting finding came from the analysis of verbal statements in studies that did not show any significant differences on problem solving tasks. One study with TAPPS demonstrated an improvement in both the amount of statements (e.g., explanations, comprehension monitoring and evaluation, and modifications to the solution of the problem) and the concepts learned (Glass, 1992). The information provided would be
useful in a school setting for teachers to monitor student use of techniques and implement correction procedures. Another study found that successful students had no negative program-assessment statements and almost half the negative self-assessment statements and not on-task statements as students that were not successful (Pate & Miller, 2011). Both of the studies that did not show any effects for TAPPS demonstrated strengths in using this technique to aid in instruction (i.e., alter instruction to reduce negative statements)

Studies in a college setting, with more advanced knowledge and required prerequisite courses performed better with TAPS with and without testing for prior knowledge (Johnson & Chung, 1999; Pate et al., 2004; Pestel, 1993). Controlling for content difficulty and choosing activities that the students have all of the prior knowledge and skills to solve may be an important part of this intervention. Despite the positive results, two of the college studies had large threats to internal and external validity that demonstrated instruction in TAPS needs more information (e.g., scoring reliability, treatment fidelity) to draw conclusions for classroom use (Johnson & Chung, 1999; Pestel, 1993).

Studies with younger students did not have the same results as the college settings. At the high school level, researchers (Jeon et al., 2005) tested the mathematical and chemistry course backgrounds for equivalence in groups, but in result may have demonstrated student prerequisite skills allowing the TAPS group to succeed in that study (i.e., like the college courses requiring prerequisites). Unlike Jeon et al.(2005), a high school study that showed no significance using TAPS did not control for content difficulty prior to the study, only demonstrating that students need to have domain specific knowledge (Pate & Miller, 2011).
Similar to the majority of high school studies, a study including middle and high school students (Glass, 1992) lacked control for content difficulty and may have contributed to students not succeeding, stating all students were naive to the concepts. Students in 7th and 8th grades may also not have had similar experience to high school or college students with problem solving and may need more preparation in content to succeed with problem solving tasks. Glass (1992) found that the statements elicited by the TAPPS group were more focused in reference to the problem solving task, but students may have been unable to solve the problem for other reasons such as lack of prior knowledge. All of the studies supporting the need for the instruction during TAPPS to be on students instructional level.

Problems with methodology raised some concerns for internal validity such as failure to report fidelity of treatment and scoring or interobserver agreement of the dependent variable for all studies. Future studies in TAPPS should report this information as well as control for the level of difficulty and practice students receive when working in pairs. Overall the studies showed benefits for TAPPS, but more information is needed for different levels of student functioning. Students with difficulty learning would also benefit from instruction solving problems (i.e., they are required to take standardized test requiring the same amount of ability to critically think).

**TAPS Research for Students with Special Needs**

TAPS is also included in some special education literature to teach students analytical thinking skills. A TAPS program has been used to teach problem solving skills, reasoning, and analytical thinking skills to students of different skill levels with and without disabilities (Robbins, 2011). In the program described, TAPS instruction included lessons teaching two
repertoires necessary to be a good speaker and listener. Students first learn the role of speaker, then once mastered, move on to the active listener role (Robbins, 2011). After learning the speaker role, the students begin to serve as their own listener. The goal of the program is a goal for science, to learn to speak effectively and to enable students to listen to themselves and others and adjust their performance. Because the students using the TAPS program may have more difficulty, students are also taught behaviors to encourage positive self-dialogue, preventing a pattern of negative statements such as those described in Pate and Miller (2011).

Ferris and Fabrizio (2009) conducted a case study with five phases testing the usefulness of TAPS for a young girl with autism to become more active in analytical thinking as a speaker. This was the only study in special education that could be located that analyzed the effect of TAPS on a specific problem solving behavior. Materials used in the study ranged from reading, math, and reasoning and writing. The goal of TAPS was to increase behaviors necessary for more critical thinking while decreasing re-reading of problems. Several important discoveries were made in the study: Two minute timings were preferable for more chances to respond; TAPS behaviors are able to be vocalized and remediated through explicit instruction and practice; and TAPS combined with a behavioral fluency component could lead to a measure for proficiency (Ferris & Fabrizio, 2009).

Another unpublished study recently conducted was Witcoski (2012), analyzing TAPS used with three students at-risk for learning disabilities in fourth grade. Science text was used in order to create problems for students to sequence sentences in the correct order. The skill of sequencing and picking out cues from reading was taught through explicit lessons and frequency building activities. Students were instructed until they reached criterion collected from high
school students advanced in science. Witcoski (2012) confirmed that a two-minute timing was preferable to allow for more problem solving behavior supporting Ferris and Fabrizio (2009) and that frequency building activities could eventually lead to a proficiency measure. All students in the study had worsening baselines prior to the intervention and all student improved in the accuracy of their problem solving performances. More information was needed to explore the intervention in depth such as maintenance data, different populations of students, generalization measures, and more replications. Witcoski (2012) demonstrated the ability to have students actively engaged in expository science text, with a problem solving and frequency building activity in fourth grade science instruction. More information is needed about applying TAPS in the science classroom.

**Current Science Instruction**

Students should begin problem solving instruction in early grades to build on prior knowledge and skills previously mastered. Science is an area that students must continue to build knowledge representing current understanding of natural systems and the process whereby that body of knowledge was established and is being continually extended, refined, and revised (Duschl, Schweingruber, & Shouse, 2007). For the process of learning to take place in science, students must actively be involved in learning, along with adapting what they already know.

Science strands were outlined in Goals of Science Education in Taking Science to School (2007) as: (a) know, use, and interpret scientific explanations of the natural world (b) generate and evaluate scientific evidence and explanations (c) understand the nature and development of scientific knowledge (d) and participate productively in scientific practices and discourse. All of the perviously listed skills need to be developed over time and should begin at the earliest age.
possible. In the United States, formal instruction begins in elementary school. Teachers in early primary school can be responsible for multiple content areas, with instruction in science beginning earlier than ever thought-in kindergarten (Duschl, Schweingruber & Schouse, 2007).

The ability to start instruction so early in a student’s schooling hinges on the amount of science knowledge a teacher has. It is known that teacher subject matter knowledge contributes to higher student achievement (Chaney, 1995; Goldhaber & Brewer, 1997, 2000). Unfortunately, teacher knowledge is often lacking in Kindergarten through 8th grades because of narrowly focused undergraduate work, insufficient professional development, and lack of credentials for this age group to teach science (Duschl, Schweingruber, & Schouse, 2007). Beginning instruction at early grades can vary widely based on what education the teacher brings to the classroom.

Instruction in science begins with the vast knowledge students come to school with based on their environments. Students continue to learn and build on prior knowledge through actively engaging in science, inquiry, and a curriculum that builds upon skills overtime with explicit connections to previously learned content (Duschl, Schweingruber, & Shouse 2007). Despite the knowledge that we have about what makes instruction better, evidence doesn’t appear to be supporting more actively engaged instruction. Less than half of American high school students report working in groups with less time devoted to student-centered discussions; with careless and poorly planned group work when it was used (Corcoran & Silander, 2009). The lack of time and quality in interacting with science content shows that when teachers are using groups, they are not effectively planned. Activities in the classroom are not entirely the fault of the teacher,
with curricular guides favoring covering many topics, rather than activities to think more deeply about science (Kesidou & Roseman, 2002).

Some improvements have been made to teacher curriculums and testing, with benchmark assessment systems being used throughout the school year. A system such as Benjamin Blooms system in which students progress through curriculum demonstrating proficiency on a set of formative assessments, has shown to have significant positive effects for lower achieving students and for inexperienced teachers (Block & Burns, 1976; Guskey & Gates, 1986; Whiting, van Burgh, and Renger, 1995). Assessing student progress is key to having knowledge to change instruction and help students. Unfortunately, the assessments in varied curriculum scored poor in providing guidance for teachers to modify their day-to-day instruction (Stern & Ahlgren, 2002). However, teachers who use specific instructional processes (i.e., receiving explicit directions about how to share information with students) to follow based on test performance and data from the assessments demonstrated significantly higher growth in student achievement than teachers who used their own judgment about how to respond to the data (Fuchs & Fuchs, 1986). All of the improvements made in science have gone in the direction of more engaging instruction with the ability to adjust lessons for students based on data.

**Conclusions**

The struggle to attain proficiency for students with special needs begins with the skills students bring to the classroom, and is complicated by current classroom practices that are incompatible with their learning needs. The large amount of knowledge along with the expectation to not only acquire but integrate thinking skills within science in problem solving activities makes learning more difficult (Kameenui & Carnine, 1998). As an intervention, TAPS
provides a guided approach that allows students to obtain immediate and relevant feedback for problem solving. Studies have shown practice talking aloud in pairs and with appropriate level of content difficulty, TAPS instruction can be beneficial for critical thinking in many content areas and science.
References


Otero, J. (2002). Noticing and fixing difficulties while understanding science texts. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (Ch. 12, pp. 281-308). Mahwah, NJ: Erlbaum.


Unpublished manuscript, Department of Special Education, The Pennsylvania State University, University Park, PA.
Appendix B

Sample of Lessons in TAPS

T=Teacher response

S= Students response

LESSON 1

Model 1

The student can make marks on the paper as they work- I will put the prompt on a separate paper, or write it on a separate sheet so he doesn’t view the lessons.

_____ T: I am going to model to you how to talk aloud when placing sentences into their most logical order. In order to be the best problem solvers, we want to talk aloud how we solve problems, in a way that others can understand. When we give the clearest explanations with a lot of detail, it lets someone like a teacher, know how we went about solving the problems. Why do you think it is important to be able to explain how we solve a problem??

_____ S: Prompt the student to give at least 2 responses (what about during class?, what about when you are working on a problem that you are unsure about one of the steps- and it has a lot of steps; How might this help you when you take a test?)

_____ T: I think that being able to talk aloud about problem solving will help you in a lot of areas in school and outside of school. You will start to understand how different people use words to connect sentences. Authors are different in that they may use different words to link two sentences, and we have to figure out the meaning based on how they presented the information. This will also help us to understand all of the different ways that we can link sentences. For example, I can show a cause and effect. An example would be: Because of poor weather, the plants were moved inside and created water stains on the floor of the house. We can also emphasize or elaborate about something we are talking about: For example; fifteen inches of snow fell overnight. We could not even get the car out of the garage! This skill should also help
you write your own sentences. By reading and evaluating how others write, you will begin to learn different ways to organize your sentences and paragraphs!

T: Read the following sentences. Number them to show their most logical order.

1. A. He planted a seedling in a pot of soil.
    B. About 400 years ago, a Dutch scientist named Jan van Helmont wanted to know how plants meet their needs.
    C. After five years, the seedling became a small tree.
    D. He watered it regularly.

A. He planted a seedling in a pot of soil. I don’t know who he is. This is most likely not the first sentence. I will wait until they introduce who he is. I also don’t know why a seedling is going to be planted in soil.

B. About 400 years ago, a Dutch scientist named Jan van Helmont wanted to know how plants meet their needs. This seems like a good introduction sentence. I found out who he is, Jan van Helmont. I also found out when this took place. Now I know why he planted a seed, because he wanted to find out this information. I will keep this one in mind for the introductory sentence.

C. After five years, the seedling became a small tree. Well, this sentence begins with after, so something had to happen before it. I am thinking that this would come after him planting it and not as the introductory sentence. You need to plant something before it grows.

D. He watered it regularly. This sentence is not an introductory sentence because we don’t know who he is or what he watered. It probably is talking about the seed. I think this should go somewhere in the middle because seeds need water to grow, so it is probably before it was a small tree.
I think the order should be, About 400 years ago, a Dutch scientist named Jan van Helmont wanted to know how plants meet their needs. He planted a seedling in a pot of soil. He watered it regularly. After five years, the seedling became a small tree.

**MODEL 2**

T: Read the following sentences. Number them to show their most logical order.

1.53.

_____ A. When the spore cases open, the spores are released.

_____ B. They drift through the air and then settle.

_____ C. The cases protect the spores from too much heat or too little water

_____ D. Spores grow inside tough spore cases

_______A. When the spore cases open, the spores are released. I do not know why spore cases open, or what they are. I should probably know something about them before I am talking about something being released from them. This probably isn’t the introduction sentence.

_______B. They drift through the air and settle. Well, I don’t know what they are, but I could guess that they might be the spores since they were released. This isn’t the first sentence because I don’t know what they are, and I think the first sentence should tell me this before I talk about them doing anything. This may be one of the last sentences, because the case needs to open first based on the first sentence.

_______C. The cases protect the spores from too much heat or too little water. I am guessing that this is a sentence about the spore cases, but that is a guess because they are released from there in the first sentence I read. This is explaining a spore case more, so I am thinking that it goes somewhere after the cases are introduced, but not first.
D. Spores grow inside tough spore cases. This is definitely the introduction sentence, because it is explaining why the spore cases and spores are mentioned in the next couple sentences. I think that this is first because it would be a good introduction to the other sentences.

I think the order should be: Spores grow inside tough spore cases. The cases protect the spores from too much heat or too little water. When the spore cases open, the spores are released. They drift through the air and settle.

Lesson 2

The student can make marks on the paper as you work- I will put the prompt on a separate paper, or write it on a separate sheet so he doesn’t see the writing below.

MODEL WITH SOME STUDENT INVOLVEMENT

We are going to work on the same type of skills today that we have focused on for the past week. Today you are going to be helping me out more because you already saw how I talked through solving a problem yesterday. Today I will ask you some questions about how to solve the problem and you can help me get the right answer.

T: Read the following sentences. Number them to show their most logical order.

1_64.

_____ A. A root pushed through the crack
_____ B. Under the right conditions, the seed will sprout
_____ C. First, the seed coat splits open.
_____ D. Then, a tiny stem grows upward.
T: This time I am going to solve this problem and ask you some questions while I am talking aloud.

T: A. What does this one say?
S: A root pushed through the crack
T: Nice reading, A root pushed through the crack. I don’t know what is happening in this paragraph yet, I am guessing that this is about a plant, and the seed grew and the root is going through some crack in the dirt or somewhere. What does B say?
S: Under the right conditions, the seed will sprout.
T: okay, so now I know that this paragraph is talking about a seed of a plant, so my first guess was right about the root. I think that this may be a good introductory sentence because it is telling us what is going to happen, but I should probably read on in case they give more information about where the seed came from. What does C say?
S: First, the seed coat splits open.
T: Well, I think there should be a sentence before this, about what is happening. That sentence may be the first one I read about the seed sprouting. This may fit because it is talking about the first step. What does the next sentence say?
S: Then a tiny stem grows upwards.
T: Great! Now I really have to think. The sentence begins with “then” so it is one of the steps that follows the first. Well, so far I read sentences about a seed, a seed coat, and a root. I would have to say that the root comes before there is a stem because the root is still in the ground. So this probably comes after the sentence about the root. And the coat has to split open around the seed before the root comes out.

I think the order is:
Under the right conditions, the seed will sprout. First, the seed coat splits open. A root pushed through the crack. Then, a tiny stem grows upward. One or two leaves appear on the stem.
MODEL with more student involvement:

T: Read the following sentences. Number them to show their most logical order. I am going to be asking you some more questions this time, so that you can help me figure out the answer to this problem. Please read the first sentence.

1. A. Others, such as seeds from the sycamore tree, have parts that look like wings.
   B. Some, like dandelion seeds, are easily carried away by wind.
   C. These “wings” help catch the wind.
   D. Seeds need space to grow, but they cannot move on their own.

S: Others, such as seeds from the sycamore tree, have parts that look like wings.

T: This sentence starts with the word others. Do you think this is a good way to start a opening sentence in a paragraph?

S: “should say no”

T: Prompt students with questions like: Do we know what “others” means? The word others is usually used to compare to something else that was previously stated. I don’t think that this goes first, because we want to know what they are comparing the seeds of a sycamore tree to. Read the next sentence.

S: Some, like dandelion seeds, are carried away by wind.

T: Okay, what words do you see in this sentence that give us an idea of where is might fall in the order?

S:

T: Prompt students with questions like: Is some a good way to start a paragraph? I know that we are talking about seeds again, so this may be the topic of the paragraph. I am guessing that this one comes before the first sentence if they aren’t talking about any other types of seeds. Read the next sentence.

S: These “wings” help catch the wind.

T: I think I heard one of these words in another sentence. Where do you think the author wanted to place this sentence?

S:
T: This sentence most likely comes after they were talking about sycamore trees having wing parts. It would make sense after that because it is describing what the wings do. Read the next sentence please.

S: D. Seeds need space to grow, but they cannot move on their own.

T: Where do you think this sentence may go and why?

S:

T: Prompt students with: Are the other sentences elaborating on this one? This seems like it is talking about all of the other sentences, so it should probably go first. It is introducing what seeds need, and how they can’t move. The other sentences go on to give examples of how some seeds will move.

T: Based on all of the details that we talked about, (write out these answers and alter based on your students discussion) we discussed that sentence D should go first. Next we said, that sentence B would be before we discussed a comparison to other types of seeds. After that sentence A would make sense and then sentence C talking about the wings. How about you read the sentences in the order we came up with and see if they fit together.

S: Seeds need space to grow, but they cannot move on their own. Some, like dandelion seeds, are easily carried away by wind. Others, such as seeds from the sycamore tree, have parts that look like wings. These “wings” help catch the wind.
Appendix C

Student Copy

MD_12
_____ A. They divide each of these groups into even smaller groups.
_____ B. Scientists divide kingdoms into smaller groups.
_____ C. They continue sorting into smaller and smaller groups.
_____ D. Each time they sort, they use the organism’s features to decide whether or not the organism belongs to the group.

MD_18
_____ A. All the animals in one group have backbones.
_____ B. Vertebrates are divided into five classes--fish, amphibians, reptiles, birds, and mammals.
_____ C. Animals in this group are called vertebrates.
_____ D. Scientists divide the animal kingdom into two main groups.

MD_24
_____ A. A newly hatched snail must find its own food.
_____ B. Depending on the temperature and the amount of moisture in the soil, the eggs “will hatch in two to four weeks.
_____ C. A young snail will first eat whatever is left of its own eggshell.
_____ D. The tiny, newly hatched snail has a fragile shell.

MD_30
_____ A. Other barriers are natural.
_____ B. Migrating amphibians often must cross busy roads as they travel to and from their breeding ponds.
_____ C. Sometimes migrating animals face barriers.
_____ D. Some barriers are made by people.
### Appendix D

Teacher Copy

MD_12
- A. They divide each of these groups into even smaller groups.
- B. Scientists divide kingdoms into smaller groups.
- C. They continue sorting into smaller and smaller groups.
- D. Each time they sort, they use the organism’s features to decide whether or not the organism belongs to the group.

Answer: B, A, C, D

Scientists divide kingdoms into smaller groups. They divide each of these groups into even smaller groups. They continue sorting into smaller and smaller groups. Each time they sort, they use the organism’s features to decide whether or not the organism belongs to the group.

MD_18
- A. All the animals in one group have backbones.
- B. Vertebrates are divided into five classes--fish, amphibians, reptiles, birds, and mammals.
- C. Animals in this group are called vertebrates.
- D. Scientists divide the animal kingdom into two main groups.

Answer: D, A, C, B

Scientists divide the animal kingdom into two main groups. All the animals in one group have backbones. Animals in this group are called vertebrates. Vertebrates are divided into five classes--fish, amphibians, reptiles, birds, and mammals.

MD_24
- A. A newly hatched snail must find its own food.
- B. Depending on the temperature and the amount of moisture in the soil, the eggs “will hatch in two to four weeks.
- C. A young snail will first eat whatever is left of its own eggshell.
- D. The tiny, newly hatched snail has a fragile shell.

Answer: B, D, A, C

Depending on the temperature and the amount of moisture in the soil, the eggs will hatch in two to four weeks. The tiny, newly hatched snail has a fragile shell. A newly hatched snail must find its own food. A young snail will first eat whatever is left of its own eggshell.

MD_30
- A. Other barriers are natural.
- B. Migrating amphibians often must cross busy roads as they travel to and from their breeding ponds.
- C. Sometimes migrating animals face barriers.

Depending on the temperature and the amount of moisture in the soil, the eggs will hatch in two to four weeks. The tiny, newly hatched snail has a fragile shell. A newly hatched snail must find its own food. A young snail will first eat whatever is left of its own eggshell.
Answer: C, D, B, A

Sometimes migrating animals face barriers. Some barriers are made by people. Migrating amphibians often must cross busy roads as they travel to and from their breeding ponds. Other barriers are natural.

MD_30
A. Migration is a natural behavior for an organism.
B. Hibernation is a state of inactivity that occurs in some animals when outside temperatures are cold.
C. Another type of natural behavior that helps an animal survive is hibernation.
D. It does not need to be learned.

Answer:
Migration is a natural behavior for an organism. It does not need to be learned. Another type of natural behavior that helps an animal survive is hibernation. Hibernation is a state of inactivity that occurs in some animals when outside temperatures are cold.

MD_33
A. Scientists dropped sweet potatoes near the monkeys’ home.
B. Scientists found that when monkeys learn new things, they teach other monkeys what they’ve learned.
C. The monkeys liked the potatoes, but they did not like the sand that stuck to the potatoes.
D. The potatoes landed in sand.

Answer: B, A, D, C
Scientists found that when monkeys learn new things, they teach other monkeys what they’ve learned. Scientists dropped sweet potatoes near the monkeys’ home. The potatoes landed in sand. The monkeys liked the potatoes, but they did not like the sand that stuck to the potatoes.
Appendix E

3/13/13
Abbie
IV 4

A. What’s hot! (supposed to the that’s)
B. Their centers may be as hot as 24,000 degrees Celsius.
C. But it’s not hot enough to heat the clouds that surround both planets. (pointed to ask words)
D. Unlike the inner planets, Jupiter and Saturn create more heat than they receive from the sun.

That’s hot.
You really don’t know what their talking about.
A pizza could be hot, anything could be hot. Except for ice cream. (4 TAP)
And then (re-read B)
Their doesn’t tell you what may be 24,000 degrees Celsius. (1 TAP) (1/2 re-read)
But D tells you about what planets and stuff is very hot. (1 TAP)
So that goes into (that’s hot) because when you would be reading this you’d think of how hot that was and you would probably think that that is really hot. (1 1/2 TAP- repeat statement, started describing how the phrase fits with another phrase) (1 re-read)

7 ½ TAP 7 ½ TAP
6 ½ correct re-read 6 ½ re reads
½ incorrect re-read ½ read
A. Some attach to the rocks along the shore as the waves slide back into the ocean.
B. The waves carry more than water.
C. All along the seashore the salty ocean waves meet the land.
D. Some tiny organisms roll in with the waves.

Um, well in this one it says some attach to rocks, we don’t know what some is. (1 1/2 TAP) 1 1/2 TAP

So we have to figure out if some is just like water. Because you know that waves is made mostly of water. (2 TAP) 2 TAP

Then the waves carry more than just water. So that’s probably going to be first because the waves do (re-read B) (2 re-read) (1 1/2 TAP) 1 1/2 TAP

And it does tell you that some things attach to the rocks that you didn’t know, but it still doesn’t tell you exactly what. (2 TAP) 2 TAP

And then (re-read C) So that one probably goes before A because the waves have to meet the land before something can get on a rock. (2 1/2 TAP) 2 1/2 TAP

And then (re-read D) so that probably, B would be first, C would go (1 TAP) 2 TAO

8 1/2 TAP 8 1/2 TAP

2 TAO 3 1/2 TAO

8 re-reads – 7 re reads
Curriculum Vita
Ginny Dembek (Witcoski)
2205 James Hall, 2900 Bedford Ave.
Brooklyn, NY 11210
(570)-401-0575 (phone)
gaw144@psu.edu (email)

Education

The Pennsylvania State University, Ph.D. in Special Education, expected December, 2013
The Pennsylvania State University, B.S. in Special Education, Certification in Elementary
Education, May, 2006

Certifications/Degrees
Pennsylvania Certified Special Education/Elementary Education Teacher
Maryland Certified Special Education Teacher
TAGteach International, Primary Certification

Teaching Experiences

Instructor of Fundamental Literacy Skills for Students with Special Needs, The
Pennsylvania State University, Spring 2012

Students with Special Needs, Brooklyn College CUNY, Fall 2013-Current

Research Methods in Special Education, Brooklyn College CUNY, Fall 2013-Current

Selected Conference Presentations

